Description of the Invention

Definitions

By "mutant polymerase" is meant a nucleic acid polymerase which has at least one altered amino acid compared to the corresponding wild-type polymerase, wherein said mutation or alteration results in the mutant polymerase having a reduced discrimination between non-canonical and canonical nucleoside triphosphates as substrates.

By "template" we mean a macromolecular pattern or mold for the synthesis of another macromolecule, composed of a sequence of nucleotides, either rNTPs or dNTPs, that serves to specify the nucleotide sequence of another structure.

"Nucleotide" refers to a base-sugar-phosphate compound. Nucleotides are the monomeric subunits of both types of nucleic acid polymers, RNA and DNA. "Nucleotide" refers to ribonucleoside triphosphates, rATP, rGTP, rUTP and rCTP, and deoxyribonucleoside triphosphates, such as dATP, dGTP, dTTP, dCTP.

As used herein, "nucleoside" refers to a base-sugar combination without a phosphate group. "Base" refers to the nitrogen-containing base, for example adenine (A), cytidine (C), guanine (G) and thymine (T) and uracil (U).

"Incorporation" refers to becoming a part of a nucleic acid polymer. There is a known flexibility in the terminology about incorporation of nucleic acid precursors. For example, the nucleotide dGTP is a deoxyribonucleoside triphosphate. Upon incorporation into DNA, it becomes a dGMP, or deoxyguanosine monophosphate moiety. Although there is no dGTP molecule in DNA, one may say that one incorporates dGTP into DNA.

As defined herein, a "canonical" nucleoside triphosphate for an RNA polymerase ("RNAP") consists of any

ribonucleoside-5'-triphosphate ("rNTP" or "NTP") which has an hydroxyl group at the 2'-position of the sugar, including, but not limited to, the four common ribose-containing substrates for an RNA polymerase -ATP, CTP, GTP and UTP. A 2'-deoxyribonucleoside-5'-triphosphate ("dNTP") which has hydrogen at the 2'-position of the sugar, including, but not limited to, the four common deoxyribose-containing substrates (dATP, dCTP, dGTP and dTTP, also known as "TTP") for a DNA polymerase ("DNAP") is defined herein as a "non-canonical" nucleoside-5'-triphosphate or a "non-canonical NTP" or a "non-canonical nucleotide" or a "non-canonical deoxynucleotide" or a "non-canonical triphosphate" or a "non-canonical substrate" for an RNA polymerase. On the other hand, a "canonical" nucleoside triphosphate for a DNAP consists of any dNTP which has a hydrogen at the 2'-position of the sugar, while an rNTP is defined as a "non-canonical NTP" or a "non-canonical nucleotide" or a "non-canonical substrate" for a DNAP. The terms "canonical" and "non-canonical" are meant to be used herein only with reference to the 2' position of the sugar. Thus, as defined herein, 2',3'-dideoxynucleoside-5'-triphosphates ("2',3'-ddNTPs" or "ddNTPs") are "non-canonical" substrates for an RNAP, but are defined as "canonical" for a DNAP. Further, any other substituent than an hydroxyl group at the 2'-position of ribose or a hydrogen at the 2'-position of deoxyribose, including, but not limited to, a fluorine ("F" or "fluoro" group) or an amino group, would be defined as "non-canonical" for both RNAPs and DNAPs herein. The terms "canonical" or "non-canonical" also are not meant to refer to the nucleic acid bases attached to the sugar moieties. Thus, for example, other natural or modified nucleic acid

bases attached to the 1'-position of ribose-5'-triphosphate would still be defined as "canonical" herein.

By "a (mutant) nucleic acid polymerase (enzyme) with reduced discrimination between canonical and non-canonical nucleoside triphosphate substrates", we have a specific quantitative definition calculated as follows:

1. One first determines the K_m and the k_{cat} for each enzyme (mutant and wild-type) using the non-canonical nucleotide as a substrate and using the canonical nucleotide as a substrate, as was described previously (Patra, et al., 1992). The value "K_m" expresses how readily the enzyme will bind the substrate (a larger K_m implies weaker binding) and "k_{cat}" expresses the rapidity with which the substrate, once bound by the enzyme, is reacted upon.
2. One next calculates the numerical value for k_{cat}/K_m for each enzyme and for each substrate. By broad scientific consensus, the specificity of an enzyme for a substrate is felt to be most suitably expressed by this ratio.
3. For each enzyme, one then calculates the numerical value obtained by using the value of k_{cat}/K_m for the canonical substrate in the numerator and k_{cat}/K_m for the non-canonical substrate in the denominator. This number indicates how much a given enzyme discriminates between the two substrates. For example, if this value equals 1, then the enzyme uses both the canonical and the non-canonical substrates equally well; it does not discriminate between the two substrates. If this value is greater than 1, then the enzyme discriminates by that factor between the two substrates; for example, if the value is 100, then the enzyme discriminates by a